

# Psychological Monographs

General and Applied

No. 444  
1957

Clifford P. Froehlich

A Criterion for Counseling

By

Clifford P. Froehlich

*University of California, Berkeley*

Price \$1.00

Vol. 71  
No. 15



Edited by Herbert S. Conrad  
Published by The American Psychological Association

# Psychological Monographs: General and Applied

*Combining the Applied Psychology Monographs and the Archives of Psychology  
with the Psychological Monographs*

HERBERT S. CONRAD, Editor

*Department of Health, Education, and Welfare  
Office of Education  
Washington 25, D.C.*

## Consulting Editors

DONALD E. BAIRD  
FRANK A. BRACH  
ROBERT G. BERNREUTER  
WILLIAM A. BROWNELL  
HAROLD E. BUKIT  
JERRY W. CARTER, JR.  
CLYDE H. COOMBS  
JOHN F. DASHUELL  
EUGENIA HANFMAN  
EDNA HEIDREYER

HAROLD E. JONES  
DONALD W. MACKINNON  
LORRIN A. RIGGS  
CARL R. ROGERS  
SAUL ROSENZWEIG  
ROSS STAGNER  
PERCEVAL M. SYMONDS  
JOSEPH TIFFIN  
LEONARD R. TUCKER  
JOSEPH ZUBIN

ARTHUR C. HOFFMAN, Managing Editor

HELEN OW, ASSOCIATE Managing Editor

Editorial Staff: FRANCIS H. CLARK, BARBARA CUMMINGS, SADIE J. DOYLE, SARAH WOMACK

## MANUSCRIPTS should be sent to the Editor:

Because of lack of space, the *Psychological Monographs* can print only the original or advanced contribution of the author. *Background and bibliographic materials must, in general, be totally excluded* or kept to an irreducible minimum. Statistical tables should be used to present only the most important of the statistical data or evidence.

Preparation of manuscripts for publication as monographs should follow the procedures given in the *APA Publication Manual*, 1957 Revision—this revision supersedes all previous manuals for preparing manuscripts for publication as monographs.

CORRESPONDENCE CONCERNING BUSINESS MATTERS (such as subscriptions and sales, change of address, etc.) should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D.C. Address changes must arrive by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee third-class forwarding postage.

COPYRIGHT, 1957, BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

## Psychological Monographs: General and Applied

A Criterion For Counseling<sup>1</sup>

CLIFFORD P. FROELICH

University of California, Berkeley

## INTRODUCTION

ONE of the most obvious objectives of counseling is to increase the counselee's knowledge of himself. This objective is justified on the grounds that a client must know himself if he is to make satisfactory adjustments to the difficulties which brought him to counseling. This justification is particularly pertinent when the problems of educational and vocational planning are considered. A client must have a reasonably accurate estimate of his mental ability *before* he can intelligently decide whether to enter college or to pursue a professional career. Likewise, he needs some knowledge of his interests *before* he can logically choose a curriculum or a vocation in harmony with them. The word "before" has been italicized in the preceding sentences to highlight a significant point in connection with this investigation which studies a criterion of counseling effectiveness. The criterion compares one perception of self with an external measure of self, and notes agreement prior to counseling. A second comparison of self-perception and external measure is made after coun-

seling, and agreement again is noted. The *criterion* is the *change in agreement between the first and second comparison*. The criterion studied combines good features of both sociological and psychological criteria. The sociological aspects of the criterion are, of course, found in the postcounseling behavior of the client. He cannot score favorably on the criterion, except by accident, unless he has increased self-knowledge. Hence, an improvement in such knowledge creates a "before" situation, subsequent to which the client is more able to show improved behavior.

The criterion in this investigation might be viewed also as a psychological one. It differs, however, from the Q sort and similar criteria used by Rogers and others in that it is based on an objective measure, a test score. Rogers (8) simply compared one perception of self with another perception of self in order to note changes. The present study compares self-perceptions obtained in the form of self-ratings with objective test data.

*Terms Used in this Study*

For the sake of brevity, certain terms will be used to denote concepts which would require fuller explanation if the following definitions were not made. The terms "congruence" and "agreement" are used interchangeably. They are used to indicate coincidence between two or more measures. "Score" represents the resulting number when the individual's

<sup>1</sup>This study was aided by a grant from the Research Fund of the Department of Education, University of California. The cooperation of Robert Brownlee, Principal, Demonstration Secondary School, made it possible to conduct the study in his school. I. Aileen Poole assisted in the collection of the data. E. Wayne Wright and Joseph Martin helped in the analysis of the data, with the use of the Computer Laboratory of the University.

actual test score is converted (in the manner to be described) to its corresponding step on a five-point scale. "Score" means, then, the quintile rank-equivalent of the test score. The word "rating" is used to indicate the numerical step on a five-point scale which the individual checked when asked to rate himself.

#### DESIGN OF STUDY AND COLLECTION OF DATA

The purpose of this study was to investigate a variable which appears to possess the quality of being a sociological-psychological criterion. To accomplish this purpose, specific questions were formulated to be answered by the data collected in the experimental situation.

1. The first of these questions was: *Is there a significant difference between the agreement of precounseling rating with score and the agreement of postcounseling rating with score?* A related question was also asked. *Are the counselees' changes in agreement in the direction of greater agreement?* In other words, do the comparisons made before counseling differ significantly from those made after counseling; and does counseling increase the congruence of self-ratings and scores? It was anticipated that if these questions were answered in the affirmative, the following one could be asked of the data.

2. *Does the counseled group at the conclusion of the experiment show more ratings which are congruent with scores than the noncounseled group?* This question formalized the expected differences between counseled and noncounseled groups of students. In contrast, the first question made explicit the idea that the number of agreements would change from before counseling to after counseling. In this case, each counselee acted as his own control. In the other, comparisons were made between groups of stu-

dents unlike in their participation in counseling.

#### Locale and Population of the Study

The basic data for this study were collected during the summer of 1953 at the Demonstration Secondary School operated by the University of California in cooperation with the Oakland, California, public schools. The summer high school completed its sixteenth session in 1953, at which time the total enrollment was 951 students. Sixty-nine per cent of the students were residents of Oakland, 30 per cent came from other California communities, and 1 per cent from outside of California. The students came from 58 high schools. Their median age was 14 years, 11 months, and their median grade placement during the preceding semester was the second semester of grade 10.

A poll of the students' reasons for attending the summer school indicates that 27 per cent were there to repeat a subject or to improve the mark, 25 per cent to gratify special interests, and the remainder were completing requirements for college entrance or high school graduation. In comparison with national norms, the student population was skewed in the direction of superior ability on a test of mental ability.

All students who during the preceding semester had been enrolled in grades nine through twelve were required to take a battery of tests during their so-called "guidance period." This period was regularly scheduled but carried no credit. The students, after testing was completed, were not expected to be at school during their guidance period.

The purpose of the testing program was explained to the students at the first testing session. They were told that they were participating in an experimental testing program which was conducted for research purposes by the University of California. They were informed also that the test results would be available to them if they wished to talk with a counselor. No effort was made to induce students to make use of the counselors. At the first testing session, students were asked whether or not they wanted to talk with a counselor. The data in Table 1 indicate the number who wished to be counseled as well as the number who were actually counseled. The discrepancy between the total of 951 students enrolled in the school and the 485 in Table 1 is accounted for by a variety of factors. Some were excused from the testing program if their out-of-school employment made it necessary, as were also foreign students whose knowledge of English was not sufficient to complete the tests. Approximately 200 students were eliminated from the study because they had not been enrolled in a high school during the preceding semester.

TABLE 1  
STUDENTS INCLUDED IN THE STUDY

Do you wish to be counseled?	Male Students		Female Students		Total	
	Counseled	Not Counseled	Counseled	Not Counseled	Counseled	Not Counseled
Yes	113	63	107	41	220	104
No	9	63	8	36	17	99
Did not answer question	3	21	5	16	8	37

Others were eliminated because their criterion or test data were incomplete. In all, 466 of the 951 students were not used in the present investigation.

### The Test Instruments

The following tests were used: *SRA Primary Mental Abilities; Intermediate (9); Kuder Preference Record-Vocational, Form C (6); Test of Mechanical Comprehension, Form BB (1); and SRA Youth Inventory, Form A (7).*

The tests of interest, mental ability, and mechanical comprehension were selected because they were judged to be typical of those used in high school counseling programs. Studies of their reliability and validity have been reported in the journal literature and publishers' manuals. For each of these tests a listing of references, as well as several critical reviews, are included in the three most recent editions of *Buros' measurement yearbooks* (2, 3, 4). All appear to be sufficiently reliable for the purposes of this investigation.

The validity of the tests used in this investigation deserves a general comment. Although they are generally accepted as being valid instruments, in one view, their validity may be considered as unrelated to the study. In this instance the study is viewed as dealing with the question of whether counseling increases the congruence between the counselee's self-rating and his score. In other words, does counseling teach a person his score so that he does reproduce it in the form of a self-rating. It is not a question of whether the score on a test of, say, outdoor interests is a valid measure of outdoor interests. The fact that the counselor accepts the score as an indication of outdoor interests and interprets it as such to the counselee lends credence to the idea that if counseling is effective counselees should show more agreement between their self-ratings and scores on tests of outdoor interests after counseling than they did before.

The criterion used in this study is concerned with the observed change when the agreement

of precounseling score and ratings is compared with the agreement of postcounseling scores and ratings. Essentially the criterion is anchored to agreement of score with rating, rather than agreement of the rating with some absolute or perfect measure of the characteristic upon which the counselee rates himself. If in counseling the score is interpreted as being indicative of a more perfect measure of the characteristic (whether or not it actually is beside the point) then the criterion should be influenced in the hypothesized manner implied in the two major questions stated previously.

The validity of the *SRA Youth Inventory* is not germane since this check list was not used as one of the variables in the criterion. It was employed, rather, in an attempt to identify differences between persons who ranked high and those who ranked low on the criterion. In a sense, an attempt was made to validate it against the criterion.

All the raw scores on the tests were converted to quintiles on the basis of the published norms with the exception of the mechanical ability test. Here no appropriate norms were available, hence quintiles were based on a sample of students making up the population used in this study.

### The Ratings

The self-ratings were secured by means of a scale designed for the study. Students were asked to rate themselves on a five-step scale for each of the ten interest areas measured by the *Kuder Preference Record*, and also their mental and mechanical ability. The interest or ability to be rated was described in nontechnical language.

Students were asked to complete the rating scale at the beginning of the first testing session. In the last week of the summer term, seven weeks after students completed the first rating scale, a sec-



and one was administered. It was identical with the first except that the instructions were changed to include a rationale for the request for repeat ratings. The second ratings were made during regular class periods.

The criterion investigated in this study was based upon these self-ratings. Such ratings have been utilized in a variety of research studies and have themselves, in some instances, been subjected to study. The literature pertinent to the topic "ratings" is voluminous; that pertaining to self-ratings is more limited. It was recognized that self-ratings had a limitation in terms of reliability. Other measures might have greater reliability but none seemed to hold as much promise for studying changes in self concept as did self-ratings. Indeed, the task of rating oneself requires one to reveal aspects of one's self concept or to resort to some sort of deception in order to conceal it. This is a compelling reason for using changes in self-estimates as a criterion of counseling. But an even more pertinent reason is to be found in the very unavailability of the ratings. One of the objectives of counseling is to eliminate or at least reduce the influence of factors which contribute to the unreliability of ratings. Counseling should improve the accuracy of the counselee's self concept. If it does, it would make it possible for him to rate himself more accurately.

At the outset of the study another limitation of the criterion was recognized: there is more involved in self-rating than just knowledge of

self. A counselee may, for example, "know" on the intellectual level that his performance on a test of mechanical comprehension is equivalent to the second centile of entering students in an engineering school. But, if he dreams of bridging the Bosphorus, irrigating the Sahara, or inventing a space ship, he may not really believe his mechanical comprehension is limited. His concept of himself as a successful engineer, figuratively speaking, insulates him from the threat of accepting the meaning of his low mechanical comprehension test score. The net result is that he rates his mechanical comprehension higher than his score rates him. The implication of this limitation for this study is that the criterion is not a measurement of changes of actual self concept but rather is assumed to be a reflection of those changes.

The reliability of the ratings used in this study was estimated from data collected in the same school situation but in a different summer. Students were asked to rate themselves and then a week later to repeat the process. No tests were given, nor were interviews held during the interval between ratings. The reliability coefficients reported in Table 2 were computed by using the customary formula for product-moment correlation and were corrected for coarse grouping in the manner described by Wert, Neidt, and Ahmann (10). All of the uncorrected coefficients are significant at the .05 level except that for "persuasive." Of the 13 corrected coefficients, 9 are significant at the .01 level. These coefficients seem to indicate that the ratings used in this study have a degree of reliability not unlike that reported in other studies of ratings. True, they

TABLE 2  
RELIABILITY COEFFICIENTS OF RATINGS

Area	N	Correlation of Ratings			
		Uncorrected		Corrected	
		Reliability Coefficient	Level of Significance	Reliability Coefficient	Level of Significance
Interests:					
Outdoor	31	.40	.05	.43	.05
Mechanical	31	.88	.01	.90	.01
Computational	31	.62	.01	.66	.01
Science	31	.39	.05	.42	.05
Persuasive	30	.25	.18	.27	.16
Artistic	31	.43	.05	.46	.01
Literary	30	.74	.01	.77	.01
Musical	31	.64	.01	.68	.01
Social Service	31	.51	.01	.51	.01
Clerical	31	.38	.05	.41	.05
Abilities:					
Mental	27	.78	.01	.81	.01
Mechanical	30	.46	.01	.50	.01

are not as high as might be desired. Nevertheless, as the writer stated earlier, ratings appear to him to have a potential for assessing self concept and this was an important factor in choosing them.

### *The Counselors*

The persons who did the counseling in the Demonstration Secondary School were all graduate students specializing in student personnel and counseling psychology in the Department of Education, University of California. At the time of their participation in this study, the counselors were enrolled in the university summer session working toward advanced degrees in counseling psychology. During the preceding semester nine had been employed as secondary school teachers, six as counselors at the college level, four as teacher-counselors in high schools, three as full-time counselors at the secondary level, two as school social workers, one as teacher-psychologist, one as head counselor, and one as guidance supervisor.

The counselors chose their own orientation to counseling. None was completely nondirective. All were aware of the fact that the pupils had been told first of the availability of counseling when they reported to the testing sessions. The promised opportunity to learn and talk over the test results was used undoubtedly by the administrators as a motivating factor during the series of five testing sessions. Hence, some pupils who reported to the counseling office came for the express purpose of test interpretation. Counselors made an effort to identify pupils who had an inadequate notion of the kind of relationship that the counselor was willing to establish. They made a special effort to make certain that all counselees realized that the counselor was willing to talk about things other than test scores.

### *The Criterion*

In counseling practice, as has been pointed out, one of the major objectives is to help the client improve his knowledge of himself. The assumption is made that for a client to make wise choices which involve him, it is necessary for

him to know himself. In some counselors' orientation, the client can best discover self-knowledge in a nondirective or client-centered relationship. Other counselors consider it a more meaningful experience if the counselor plays a more active role. Although counselors vary greatly in their relative adherence to either of these two points of view, there is rather general agreement among them that an important outcome of counseling is an increase in the client's knowledge of himself. The criterion investigated in this study was designed to measure this outcome.

The criterion was change in agreement between a person's rating of his ability or interest and an objective measurement thereof. In short, the client's ratings on a five-point scale in each area were compared with quintile ranks based on his test scores in corresponding areas. In terms of the desired outcome—improvement in knowledge of self—certain assumptions had to be made concerning the criterion data.

Although it was recognized that the test scores were fallible and perhaps might not be accurate, it was assumed that they did reflect an image of the client. From this assumption it followed that if the client had accurate knowledge of self his view of himself would correspond with the test-revealed one. Conversely, it was expected, therefore, that clients would learn about themselves from test scores.

A second assumption was made that if a person knew what his tested abilities and interests were he would be able to rate them accurately. Conversely, if he did not know what his abilities were, then he would be unable to rate them accurately.

The third assumption was made that although the subjects of this study would undoubtedly have a considerable amount of information about themselves, the process of test-taking and counseling would improve this knowledge. These three assumptions were basic to the design of the study.

The data are presented for three groups of students: *Counseled*, *counseled-without-test*, and *not counseled*.

In the *counseled* group were included those students who were counseled and with whom the counselor discussed the results of the test. It should be noted that the counseled group varied from test to test; that is, not all persons who were counseled had all test results interpreted for them. One counselee, for example, may have had the interest inventory and mental ability test interpreted while another may have discussed the interest inventory and the mechanical aptitude test with the counselor. Both counselees would be included in the counseled group for the ten scores of the interest inventory. The first counselee would be included in the counseled group for the mental ability test and in counseled-without-test group for the mechanical aptitude test. Conversely, the second would be included in the counseled group of the mechanical aptitude test and counseled-without-test group on the mental ability test.

The second group, *counseled-without-test*, includes those persons who were counseled but for whom the particular test was not interpreted. Noninterpretation of tests was occasioned by several factors such as a delay in scoring the test so that it was not available at the time of counseling or the judgment of the counselor that the test score would not make a significant contribution to the counseling.

The third group, *"not counseled,"* includes those individuals who took the tests, but who were not counseled. Counseling was available for all who desired it but no effort was made to induce students to seek counseling. Of the 116 students who indicated at the first testing session that they did not wish to be counseled, only 17 later changed their minds and sought counseling.

#### FINDINGS

The data concerning each test were separated into three groups according to whether or not the individual was counseled, counseled-without-test, or not counseled. In order to study the agreement between scores and ratings, each individual's score and corresponding rating were compared. If both fell on the same numerical step of the five-point scale, it was tallied as agreement. That is, if the individual's score was a "five" and his rating a "five," the values were considered in agreement. But if his score was "five" and his rating "four," the values were tallied as disagreement. The agree-

ment between first rating and score was determined, as was also the agreement between second rating and score. Each individual on the basis of agreement between score and ratings of each area was classified into one of four categories, namely:

*Category 1.* The score and first rating disagreed and so did the score and second rating. Persons in this category consistently rated themselves inaccurately.

*Category 2.* The score and first rating agreed and so did the score and second rating. These individuals were consistent in rating themselves accurately.

*Category 3.* The score and first rating disagreed, but the score and the second rating agreed. Included in this group are those individuals whose first rating was incorrect but who correctly rated themselves on the second rating. They changed from "disagree" to "agree."

*Category 4.* The score and first rating agreed but the score and second rating did not. These persons changed their rating from an accurate one to an inaccurate one. They went from "agree" to "disagree."

Percentages based on the data tabulated according to the above-described categories are shown in Table 3. It is obvious that the majority of individuals' ratings consistently disagreed with their score. This does not mean that they did not change their ratings. They might have rated themselves too low on the first rating and too high on the second; in both instances score and rating disagreed. The median percentage for this category for the counseled group was 55, for the counseled-without-test group was 54, and for the not-counseled group was 61.

These median percentages are indicative of a situation of vital concern to counselors. Why is it that the majority of pupils in this study were unable to rate accurately their abilities on even such a gross scale as the five-point one used here? In a preceding section it was pointed out



TABLE 3  
PERCENTAGES FOR CATEGORIES OF AGREEMENT BY GROUPS AND AREAS

Area	Counseled					Counseled-Without-Test					Not-Counseled				
	N	Consistently Disagree	Consistently Agree	Disagree to Agree	Agree to Disagree	N	Consistently Disagree	Consistently Agree	Disagree to Agree	Agree to Disagree	N	Consistently Disagree	Consistently Agree	Disagree to Agree	Agree to Disagree
Interests:															
Outdoor	175	69	12	13	06	66	64	11	11	14	229	72	08	10	10
Mechanical	174	55	17	13	15	66	64	20	08	08	230	70	11	11	08
Computational	175	50	15	17	18	66	43	26	17	14	229	60	14	14	16
Science	175	47	20	19	14	66	57	15	11	17	230	56	18	11	15
Persuasive	175	53	18	20	09	66	54	14	18	14	229	63	12	13	12
Artistic	175	58	20	11	11	66	46	17	20	17	230	47	19	10	18
Literary	175	52	22	15	11	66	54	23	14	09	229	59	14	14	13
Musical	175	53	21	15	11	66	47	24	15	14	230	51	20	12	11
Social Service	175	55	17	19	09	66	45	18	17	20	229	63	11	12	14
Clerical	175	56	17	12	15	66	48	14	23	15	230	61	16	12	11
Abilities:															
Mental	150	68	15	11	06	75	66	17	09	08	218	66	11	10	13
Mechanical	79	60	14	15	11	158	62	13	20	15	229	61	13	14	11
Musical	99	57	19	10	14	128	67	10	09	08	212	64	10	08	12

that self-ratings were used because they required the subject to reveal aspects of his self concept or resort to deception. The data in this study cannot yield the reason for the inaccuracies. They do, however, hint at the possibility that lack of agreement between rating and test score was brought about not so much by a lack of "knowledge" of self as it was by a lack of "acceptance" of the knowledge or an unwillingness to reveal the knowledge through ratings. Consider, for example, the rating of mental ability. All of the subjects in the study had had extensive experience in school which provided many opportunities for them to acquire a sound basis for judging their mental ability. School marks, teacher statements, feelings of success and failure, parent judgments, and peer evaluations are illustrative of the experiences which would help a person acquire an accurate self concept. Prior to counseling, 79 per cent of the subjects inaccurately rated their mental ability. After counseling, when they would have been given the knowledge necessary to rate themselves accurately, 68 per cent persisted in rating themselves inaccurately.

#### Comparison of Interest Profiles

The *Kuder Preference Record—Vocational, Form C* yields ten scores, each indicative of interests in a particular area. Counselors usually interpret individual scores in the light of the total profile.

Thus, one student's highest scores might be at the 85th centile in the clerical and musical areas, while another student's highest scores are in the same area but are equivalent to centiles of

60. In both cases, the counselor would probably interpret the highest scores as indicative of areas of greatest interest. But the interpretation in both would continue to the point of helping the client see the strength of his interests in relation to those of persons in the norm group. Essentially, profile interpretation is a process of discerning the strengths of interest in the ten areas measured by the Kuder in relation to each other and to the normative population. Because such interpretations were common practice among the counselors included in this study it appeared desirable to consider the accuracy with which students reported their obtained Kuder scores when taken together as a profile. To make the profile comparisons, the data were reprocessed using a technique of assessing similarity between profiles (5). The technique is illustrated by the following data concerning one student:

Kuder Interest Areas										
Column	0	1	2	3	4	5	6	7	8	9
First rating	5	2	5	2	4	4	4	5	4	1
Test score	3	4	5	2	5	2	5	3	4	1
Difference	2	-2	0	0	-1	2	-1	2	0	0
Square of difference	4	4	0	0	1	4	1	4	0	0

In the "0" (Outdoor Interest) column, the "5" is the student's rating of his outdoor interest which he made before taking the test. The "3" is the quintile category based on his obtained test score. The "2" is the difference when the score is subtracted from the rating. And finally, the "4" is the square of the difference between score and rating. The data for the other areas were treated in a similar manner. A figure

of 18 is obtained by adding the numbers in bottom row above. The square root of 18, or 4.243, was used as this student's first-rating-profile score.

Another profile score was obtained for this student by comparing the interest ratings he made after counseling with his test scores. The technique used followed the pattern described above. This process yielded a second-rating-profile score. In the case of the student cited above, his second-rating-profile score was 3.464. Similarly, the data concerning all other students in the study were processed to yield for each student a first-rating-profile score and a second-rating-profile score.

Then the second-rating-profile score was subtracted from the first-rating-profile score. In the case of the student used as an example in the preceding paragraphs this computation yielded a positive number, .779 (i.e.,  $4.243 - 3.464 = .779$ ). The algebraic difference between the profile scores indicated whether the individual had moved toward greater congruency between the profile of ratings and the profile of scores, or toward less congruence. A positive difference between the profile scores indicated that the second profile of ratings corresponded more nearly to the obtained profile of scores than did the first profile of ratings. Conversely, a negative difference means that the first ratings were more consonant with scores than were second ratings. In fine, positive differences showed increased "accuracy" and negative differences a deterioration in "accuracy."

Separate frequency distributions of the differences between profile scores were made for the counseled, counseled-without-test, and not-counseled groups. The distributions ranged from high positive through zero difference to high negative.

The zero-difference category was not interpreted as indicating that the student did not change his profile of ratings. The zero difference merely indicated that the subject's second profile of ratings agreed neither less nor more closely with the profile of scores than did his first rating profile.

For the distributions of the differences between the first-rating-profile score and the second-rating-profile score the means and standard deviations were computed. Table 4 contains these data. The means

TABLE 4  
MEAN AND STANDARD DEVIATION OF  
DIFFERENCES BETWEEN PROFILE SCORES

Statistic	Coun- seled	Coun- seled With- out Test	Not Coun- seled
Mean of difference	.41	.26	.04
Standard error of mean	.09	.12	.07
Standard deviation	1.09	.95	.97
Number of cases	175	66	230

varied in magnitude in the expected manner; namely, the counseled group showed the most change in the direction of greater agreement, the not-counseled group the least change. A *t* test indicates that the means of the counseled and the counseled-without-test groups are significantly different from zero. In considering the mean difference of .41, one should not overlook the fact that the difference between score and rating could not exceed 4.00. Hence the difference is slightly more than 10 per cent of the largest possible difference. To test the significance of the difference between means, *t* ratios were computed. The *t* ratio between the mean of the counseled group and the mean of the noncounseled group was 3.52, significant at the .01 level. The *t* ratios between the mean of

TABLE 5  
COMPARISON OF CERTAIN CHARACTERISTICS OF SUBGROUPS SELECTED ON BASIS OF ACCURACY  
OF RATING OF INTEREST

Variable	Counseled		Counseled Without Interest Test		Not Counseled	
	Increased	Deteriorated	Increased	Deteriorated	Increased	Deteriorated
Mean age	15.8	15.9	16.1	16.3	16.1	15.7
Mean school grade	10.4	10.6	10.7	10.6	10.3	10.3
Mean P.M.A. score	3.3	3.5	3.1	2.5	3.1	3.1
Mean number of tests interpreted	5.0	4.7	3.7	3.5	—	—
Number of: males	23	24	11	7	36	38
females	24	23	7	11	26	24
Number: desired counseling	45	45	15	12 <sup>a</sup>	25 <sup>a</sup>	30 <sup>a</sup>
did not desire counseling	2	2	3	3 <sup>a</sup>	28 <sup>a</sup>	23 <sup>a</sup>

\* Some students did not respond to the question, "Would you like an opportunity to talk with a counselor?" Therefore, these totals differ from others in the table. In instances where the number of nonrespondents was excessive, tests of significance were not applied.

the counseled-without-test group and the counseled and the not-counseled groups were found to be 1.03 and 1.63, respectively. Neither was statistically significant.

The data thus far reported have answered affirmatively the two questions spelled out in the section which dealt with the purpose of the study. They revealed that there was a significant difference between precounseling and post-counseling criterion measurements. The data also revealed that these observed differences were in the direction of more agreements after counseling. The stability of these differences was also demonstrated by the differences between the counseled, counseled-without-test, and not-counseled groups. The neat stair-stepping of mean differences in Table 4 from smallest for the non-counseled group to largest for the counseled group lends credence to the belief that the observed changes are in part at least a function of counseling.

#### *Comparison of Factors*

In this section data will be presented

concerning factors thought to be associated with changes measured by the criterion. They are: age, grade placement, intelligence, sex, number of tests interpreted, and desire for counseling.

In order to determine if the factors were associated with the criterion, the subjects were divided on the basis of the difference between the first-rating-profile score and the second-rating-profile score. The 27 per cent of the students with the largest positive differences (increase in "accuracy") and the 27 per cent of the students with the largest negative differences (deterioration in "accuracy") in the area of interests were identified in the counseled, counseled-without-test, and not-counseled groups. This process yielded the six subgroups.

For each subgroup the statistics summarized in Table 5 were obtained. When appropriate tests of significance were applied to differences between the "increased" and the "deteriorated" subgroups in the counseled, counseled-without-tests, and not-counseled groups, none of the differences was found to be statistically significant. Comparable results

TABLE 6  
NUMBER OF PROBLEMS CHECKED ON THE *SRA Youth Inventory*

Inventory Areas	Counseled					Counseled Without Test					Not Counseled				
	Increased N=34		Deteriorated N=29		t ratio	Increased N=14		Deteriorated N=12		t ratio	Increased N=55		Deteriorated N=55		t ratio
	Mean	SD	Mean	SD		Mean	SD	Mean	SD		Mean	SD	Mean	SD	
My School	6.00	4.34	5.31	3.48	.68	6.64	3.87	4.92	3.17	1.20	7.25	4.40	6.44	4.50	-.94
After High															
School	11.68	7.67	8.97	6.09	1.54	13.14	7.56	9.25	7.43	1.27	9.96	7.65	9.02	6.77	.68
About Myself	6.70	6.54	6.14	4.89	.38	8.86	5.29	6.75	8.70	.70	7.89	7.84	6.45	5.87	1.07
Getting Along															
with Others	6.79	6.42	6.93	4.50	.10	8.29	6.17	4.67	5.02	1.59	6.96	6.36	6.22	6.22	.61
My Home and															
Family	4.18	6.12	3.52	5.35	.45	4.36	3.29	2.50	3.23	1.05	5.11	6.46	3.96	5.85	-.07
Boy Meets Girl	3.91	3.93	3.93	4.30	.02	3.57	3.81	2.92	1.89	.54	3.56	4.77	3.24	3.55	.40
Health	2.62	2.29	2.79	2.59	.27	2.57	2.82	2.33	1.97	.25	3.05	3.14	2.65	2.32	.74
Things in															
General	4.18	4.34	3.66	3.17	.54	4.21	4.99	4.00	3.51	.38	3.93	4.37	3.55	4.74	.44
Basic Difficulty															
Score	12.41	11.24	11.66	8.58	.30	14.14	7.94	12.50	13.50	.35	15.05	12.76	12.65	10.20	1.08

were found when mental ability and mechanical aptitude were similarly treated.

On the basis of these data, it was concluded that the improvement or deterioration in the "accuracy" of students' self-ratings apparently was not a function of their age, school grade, mental ability, sex, desire for counseling, or number of tests interpreted during counseling. It appeared, therefore, that the criterion employed in this study will be useful in the further study of counseling because it was not influenced by gross factors over which counselors have little control. This conclusion led to a desire to learn even more about factors associated with the criterion. One attempt to study such factors will be reported in the next section.

#### Comparison of Problems

The *SRA Youth Inventory* appeared to be an instrument which might be used to identify some of the more subtle characteristics which distinguished those students who show more agreements between scores and ratings after counseling than before. The *Inventory* is a list of 298 problems of which the testee checks those which apply to him. The more problems he checks the higher his score.

Did students who showed more agreements have more or fewer problems than those who showed fewer agreements after counseling? The answer to this question was sought in a comparison of the *SRA Youth Inventory* scores of 100 students who showed the greatest increase in agreements with 99 students who showed the greatest deterioration in the agreement between their profiles of interest scores and ratings. These students were drawn from the upper and lower 27 per cents of the distribution of differences between first- and second-rating-profile scores previously described. The mean inventory score of the "increased" group was 46.92 with a standard deviation of 30.32. The "deteriorated" group had a mean inventory score of 41.72; the standard deviation was 29.64. The *t* ratio for the difference between these means was 1.22, not significant. Hence, the question asked above was answered: there was no difference between the groups in terms of the total number of problems checked on the *Inventory*.

The next step was to determine whether or not there was a difference between the "increased" and "deteriorated" groups in the number of problems checked in specific areas. The *SRA Youth Inventory*, as previously pointed out, yielded scores in eight areas of adjustment. In order to further refine the analysis, the groups were broken down into counseled, counseled-without-test, and not-counseled subgroups. The mean number of problems checked by each subgroup for each area is presented in Table 6. In Table 6 the *t* ratios based on the differences between means are also given. None was found to be statistically significant. Hence, it was concluded that the *SRA Youth Inventory* scores did not identify those students who showed more agreements between their profile of scores and ratings after counseling than before.

The two comparisons based on number of problems checked did not take into account the

possibility that certain particular problems might be checked by the "increased" group and others by the "deteriorated" group. The previously discussed analyses were based simply upon number of problems checked. Hence an item analysis was undertaken. The groups were divided into random halves. The resulting groups were (a) Increased I,  $N=50$ ; (b) Increased II,  $N=50$ ; (c) Deteriorated I,  $N=49$ ; and (d) Deteriorated II,  $N=50$ . The item responses of the Increased I and Deteriorated I groups were tabulated and percentage of response for each group for each item was computed. If the difference between the percentages for the two groups was significant at the .05 level, the item was retained.

There were 35 of the 298 items in the test retained. Of these, 32 were keyed for increase and 3 keyed for deterioration. These items were cast into a scoring stencil and all answer sheets for the test were rescored. A single score was obtained by subtracting the deterioration-keyed items from the increase-keyed items. The mean scores and standard deviations were computed. These data are presented in Table 7. The differences between the means of the Increased I and Deteriorated I groups yield a  $t$  ratio of 7.58, highly significant. This was expected since the scoring key was based on these groups. In an attempt to cross validate, the 35-item scoring key was applied to the answer sheets of the subjects included in the Increased II and Deteriorated II groups. The data in Table 7 con-

tion of the number or kind of problems the student checked.

#### SUMMARY AND DISCUSSION

The over-all purpose of this investigation was to study a proposed criterion for the evaluation of counseling, namely, *change in agreement between self-rating and score*. It is appropriate to ask, "What was learned about the criterion?" Details in the answer to this question can be found in the preceding pages. They may be summarized briefly. The criterion data varied in the expected direction, that is, the findings followed the logic of improved self-knowledge after counseling. From this point of view the criterion which was adopted appears to hold promise as a useful indicator of counseling effectiveness.

Its potential usefulness was further substantiated by the findings that the criterion appeared not to be affected by age, sex, school grade placement, intelligence, or desire for counseling. These are factors over which a counselor has no control (except by limiting his practice to specified types of clients). The important implication of these findings is that the criterion variable is independent of such factors which are extraneous to the counseling process *per se*.

A related finding was that neither the number nor the kind of "problem" which counselees verbalized on the problem checklist was associated with the criterion variable. If this result were accepted at face value, it would do violence to beliefs which counselors have about their relative effectiveness with clients having "minor" as opposed to "major" problems of adjustment, or with clients having certain kinds of problems (e.g., vocational choice as opposed to feelings of inferiority). In considering these data, it is important to

TABLE 7  
MEAN AND STANDARD DEVIATIONS ON SCORES  
DERIVED FROM ITEM-ANALYSIS KEY  
(Cross-Validation Sample)

Group	N	Mean	Standard Deviations
Increased I	50	8.00	4.56
Deteriorated I	50	2.51	2.20
Increased II	50	4.76	2.84
Deteriorated II	49	5.18	3.54

cerning these two groups indicated that the 35-item key did not hold up. When a  $t$  test was applied to the difference between the means of these two groups, it yielded a  $t$  ratio of .65, not significant.

These analyses of the *SRA Youth Inventory* data suggested that the increased-deteriorated classification was not a func-



keep in mind that the criterion involved ratings in the areas of interests and ability. It was not concerned with so-called problems of personal adjustment, the forte of the problem checklist used in this investigation. But even in areas of the problem checklist closely related to interests and ability, no statistically significant association with the criterion was noted.

This brief synopsis of the findings of this investigation can serve as a starting point for considering next steps.

The criterion explored in this investigation is an "immediate" one, i.e., it is available within a short time after the experiment is completed. But its relationship to the long-range objectives of counseling has not been established. Herein lies a great weakness. If one were to argue that self-knowledge is the object of counseling, the criterion studied might be presumed to be a reasonable measure. However, few counselors, if any, would be satisfied with self-knowledge as an end in itself. The common expectation is that

self-knowledge should influence counselee behavior. Here the criterion is silent, for it does not measure behavior. And in this silence is found an important gap in the usefulness of the criterion, as well as a suggestion for additional research. In future research it might be possible, for example, to discover that the criterion is associated with more intelligent, more constructive, or more adaptive postcounseling behavior. If such a discovery were made, then the criterion would acquire sociological significance.

Another fact must be considered. Some students who were not counseled showed as much change on the criterion in the direction of more "accurate" ratings as did students who were counseled. Conversely, some counselees deteriorated in the "accuracy" of their ratings as much as some not-counseled students. What factors are associated with change as measured by the criterion? The design of the present experiment was such that only counseling as total process was studied. Are there variations within the process from counselor to counselor, counselee to counselee, or problem to problem which account for the observed differences in the criterion variable? If so, are these factors also at work in different form among not-counseled students? These are crucial questions for which no ready answers are available.

## REFERENCES

1. BENNETT, G. K., & FRY, DINAH E. *Test of mechanical comprehension, Form BB*. New York: Psychological Corp., 1941.
2. BUROS, O. K. (Ed.) *The nineteen forty mental measurements yearbook*. Highland Park, N.J.: The Mental Measurements Yearbook, 1941.
3. BUROS, O. K. (Ed.) *The third mental measurements yearbook*. New Brunswick, N.J.: Rutgers Univer. Press, 1949.
4. BUROS, O. K. (Ed.) *The fourth mental measurements yearbook*. Highland Park, N.J.: Gryphon Press, 1953.
5. CRONBACH, L. J., & GLESER, GOLDINE C. Assessing similarity between profiles. *Psychol. Bull.*, 1953, 50, 456-473.
6. KUDER, G. F. *Kuder preference record-vocational, Form C*. Chicago: Science Research Associates, 1951.
7. REMMERS, H. H., & SHIMBERG, B. *SRA youth inventory, Form A*. Chicago: Science Research Associates, 1949.
8. ROGERS, C. R., & DYMOND, ROSALIND F. (Eds.) *Psychotherapy and personality change*. Chicago: Univer. of Chicago Press, 1954.
9. THURSTONE, L. L., & THURSTONE, THELMA G. *SRA primary mental abilities: intermediate*. Chicago: Science Research Associates, 1949.
10. WERT, J. E., NEIDT, C. O., & AHMANN, J. S. *Statistical methods in educational and psychological research*. New York: Appleton-Century-Crofts, 1954.

(Accepted for publication March 11, 1957)



GEORGE BANTA COMPANY, INC., NEW YORK